09/489,667

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12N 15/31, 1/21, C12P 21/02, C07K 14/33, A61K 38/16, 39/08

A1

(11) International Publication Number: WO 98/078

(43) International Publication Date: 26 February 1998 (26.02.9

(21) International Application Number: PCT/GB97/02273

(22) International Filing Date: 22 August 1997 (22.08.97)

(30) Priority Data:
9617671.4 23 August 1996 (23.08.96) GB
9625996.5 13 December 1996 (13.12.96) GB

(60) Parent Application or Grant
(63) Related by Continuation
US 08/782,893 (CIP)
Filed on 27 December 1996 (27.12.96)

(71) Applicants (for all designated States except US): MICROBIOLOGICAL RESEARCH AUTHORITY CAMR (CENTRE FOR APPLIED MICROBIOLOGY & RESEARCH) [GB/GB]; Porton Down, Salisbury, Wiltshire SP4 0JG (GB). THE SPEYWOOD LABORATORY LIMITED [GB/GB]; 14 Kensington Square, London W8 5HH (GB).

(72) Inventors; and
(75) Inventors/Applicants (for US only): SHONE, Clifford, Charles [GB/GB]; Microbiological Research Authority CAMR (Centre for Applied Microbiology & Research), Porton Down, Salisbury, Wiltshire SP4 0JG (GB). QUINN, Conrad, Padraig [GB/GB]; Microbi logical Resear Authority CAMR (Centre for Applied Microbiology & R search), Porton Down, Salisbury, Wiltshire SP4 0JG (GB FOSTER, Keith, Alan [GB/GB]; Microbiological Resear Authority CAMR (Centre for Applied Microbiology Research), Porton Down, Salisbury, Wiltshire SP4 0J (GB).

(74) Agent: SCHLICH, George, William; Mathys & Squire, 10 Gray's Inn Road, London WC1X 8AL (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BF BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE HU, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT RO, RU, SD, SE, SG, SI, SK, TJ, TM, TR, TT, UA, UG US, UZ, VN, ARIPO patent (GH, KE, LS, MW, SD, SZ UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI paten (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD TG).

Published

With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: RECOMBINANT TOXIN FRAGMENTS

(57) Abstract

A polypeptide has first and second domains which enable the polypeptide to be translocated into a target cell or which increase the solubility of the polypeptide, or both, and further enable the polypeptide to cleave one or more vesicle or plasma-membrane associated proteins essential to exocytosis. The polypeptide thus combines useful properties of a clostridial toxin, such as a botulinum or tetanus toxin, without the toxicity associated with the natural molecule. The polypeptide can also contain a third domain that targets it to a specific cell, rendering the polypeptide useful in inhibition of exocytosis in target cells. Fusion proteins comprising the polypeptide, nucleic acids encoding the polypeptide and methods of making the polypeptide are also provided. Controlled activation of the polypeptide is possible and the polypeptide can be incorporated into vaccines and toxin assays.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece	ML	Mali	TR	Turkey
BG	Bulgaria	HU	Hungary	MN	Mongolia	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MR	Mauritania	UA	Ukraine
BR	Brazil	IL	Israel	MW	Malawi	UG	Uganda
BY	Belarus	IS	Iceland	MX	Mexico	US	United States of America
CA	Canada	IT	Italy	NE	Niger	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NL	Netherlands	VN	Viet Nam
CG	Congo	KE	Kenya	NO	Norway	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NZ	New Zealand	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	PL	Poland		
CM	Cameroon	KR	Republic of Korea	PT	Portugal		
CN	China	KZ	Kazakstan	RO	Romania		
CU	Cuba	LC	Saint Lucia	RU	Russian Federation		
CZ	Czech Republic	LI	Liechtenstein	SD	Sudan		
DE	Germany	LK	Sri Lanka	SE	Sweden		
DK	Denmark	LR	Liberia	SG	Singapore		
EE	Estonia						

RECOMBINANT TOXIN FRAGMENTS

This invention relates to recombinant toxin fragments, to DNA encoding thes fragments and to their uses such as in a vaccine and for *in vitro* and *in vivo* purposes.

The clostridial neurotoxins are potent inhibitors of calcium-depend nt neurotransmitter secretion in neuronal cells. They are currently consider d to mediate this activity through a specific endoproteolytic cleavage of at least on f three vesicle or pre-synaptic membrane associated proteins VAMP, syntaxin r SNAP-25 which are central to the vesicle docking and membrane fusion events of neurotransmitter secretion. The neuronal cell targeting of tetanus and botulinum neurotoxins is considered to be a receptor mediated event following which th toxins become internalised and subsequently traffic to the appropriate intrac llular compartment where they effect their endopeptidase activity.

The clostridial neurotoxins share a common architecture of a catalytic L-chain (LC, ca 50 kDa) disulphide linked to a receptor binding and translocating H-chain (HC, ca 100 kDa). The HC polypeptide is considered to comprise all or part of two distinct functional domains. The carboxy-terminal half of the HC (ca 50 kDa), termed the H_C domain, is involved in the high affinity, neurospecific binding of th neurotoxin to cell surface receptors on the target neuron, whilst the amino-terminal half, termed the H_N domain (ca 50 kDa), is considered to mediate the translocati n of at least some portion of the neurotoxin across cellular membranes such that th functional activity of the LC is expressed within the target cell. The H_N domain als has the property, under conditions of low pH, of forming ion-permeable chann ls in lipid membranes, this may in some manner relate to its translocation functi n.

For botulinum neurotoxin type A (BoNT/A) these domains are considered to r side within amin acid residu s 872-1296 f r the H_C, amin acid r sidu s 449-871 for th H_N and r sidu s 1-448 f r the LC. Dig stion with trypsin eff ctiv ly d grad s the H_C d main of th BoNT/A t generat a non-t xic fragment designat d LH_N,

which is no longer able to bind to and ent r neurons (Fig. 1). The LH_N fragm nt so produced also has the property of enhanced solubility compared to both the par nt holotoxin and the isolated LC.

It is therefore possible to provide functional definitions of the domains within th neurotoxin molecule, as follows:

(A) clostridial neurotoxin light chain:

-a metalloprotease exhibiting high substrate specificity for vesicle and/or plasma - membrane associated proteins involved in the exocytotic process. In particular, it cleaves one or more of SNAP-25, VAMP (synaptobrevin / cellubrevin) and syntaxin:

(B) clostridial neurotoxin heavy chain H_N domain:

-a portion of the heavy chain which enables translocation of that portion of the neurotoxin molecule such that a functional expression of light chain activity oc urs within a target cell.

-the domain responsible for translocation of the endopeptidase activity, following binding of neurotoxin to its specific cell surface receptor via the binding domain, into the target cell.

-the domain responsible for formation of ion-permeable pores in lipid membranes under conditions of low pH.

-the domain responsible for increasing the solubility of the entire polypeptid compared to the solubility of light chain alone.

(C) clostridial n ur toxin heavy chain H_C d main.

-a porti n f th heavy chain which is r sponsible for binding f th native

holotoxin to cell surface receptor(s) involv d in th intoxicating acti n f clostridial toxin prior to internalisation of the toxin into the cell.

The identity of the cellular recognition markers for these toxins is currently not understood and no specific receptor species have yet been identified although Kozaki et al. have reported that synaptotagmin may be the receptor for botulinum neurotoxin type B. It is probable that each of the neurotoxins has a diff rent receptor.

It is desirable to have positive controls for toxin assays, to develop clostridial toxin vaccines and to develop therapeutic agents incorporating desirable properties of clostridial toxin.

However, due to its extreme toxicity, the handling of native toxin is hazard us.

The present invention seeks to overcome or at least ameliorate problems associated with production and handling of clostridial toxin.

Accordingly, the invention provides a polypeptide comprising first and s cond domains, wherein said first domain is adapted to cleave one or more vesicl r plasma-membrane associated proteins essential to neuronal exocytosis and wh r in said second domain is adapted (i) to translocate the polypeptide into the cell or (ii) to increase the solubility of the polypeptide compared to the solubility of th first domain on its own or (iii) both to translocate the polypeptide into the cell and to increase the solubility of the polypeptide compared to the solubility of the first domain on its own, said polypeptide being free of clostridial neurotoxin and fr of any clostridial neurotoxin precursor that can be converted into toxin by prote lytic action. Accordingly, the invention may thus provide a single polypeptide hain containing a domain equivalent to a clostridial toxin light chain and a domain providing th functi nal asp cts f th H_N f a cl stridial t xin heavy chain, whilst lacking th functi nal asp cts f a cl stridial t xin H_C d main.

For the purposes of the inv ntion, the functional property or properties of the H_N of a clostridial toxin heavy chain that are required to be exhibited by the s cond domain of the polypeptide of the invention are either (i) translocation of the polypeptide into a cell, or (ii) increasing solubility of the polypeptide compared to solubility of the first domain on its own or (iii) both (i) and (ii). References hereafter to a H_N domain or to the functions of a H_N domain are references to this property or properties. The second domain is not required to exhibit other properties of th H_N domain of a clostridial toxin heavy chain.

A polypeptide of the invention can thus be soluble but lack the translocation function of a native toxin-this is of use in providing an immunogen for vaccinating or assisting to vaccinate an individual against challenge by toxin. In a sp cific embodiment of the invention described in an example below a polypeptide designated LH_{423}/A elicited neutralising antibodies against type A neurotoxin. A polypeptide of the invention can likewise thus be relatively insoluble but retain the translocation function of a native toxin - this is of use if solubility is imparted to a composition made up of that polypeptide and one or more other components by one or more of said other components.

The first domain of the polypeptide of the invention cleaves one or more vesicle r plasma-membrane associated proteins essential to the specific cellular process f exocytosis, and cleavage of these proteins results in inhibition of exocyt sis, typically in a non-cytotoxic manner. The cell or cells affected are not restricted t a particular type or subgroup but can include both neuronal and non-neuronal c lls. The activity of clostridial neurotoxins in inhibiting exocytosis has, indeed, been observed almost universally in eukaryotic cells expressing a relevant cell surfac receptor, including such diverse cells as from Aplysia (sea slug), Drosophila (fruit fly) and mammalian nerve cells, and the activity of the first domain is t b understood as including a corresponding range of cells.

Th polyp ptid of th invention may b obtain d by xpression of a recombinant nucleic acid, preferably a DNA, and is a singl polyp ptid , that is t say not

cleav d into separate light and heavy chain domains. Th polyp ptide is thus availabl in convenient and large quantities using recombinant techniques.

In a polypeptide according to the invention, said first domain preferably comprises a clostridial toxin light chain or a fragment or variant of a clostridial toxin light chain. The fragment is optionally an N-terminal, or C-terminal fragment of th light chain, or is an internal fragment, so long as it substantially retains the ability to cleave the vesicle or plasma-membrane associated protein essential to exocytosis. The minimal domains necessary for the activity of the light chain of clostridial toxins are described in J. Biol. Chem., Vol.267, No. 21, July 1992, pages 14721-14729. The variant has a different peptide sequence from the light chain or from the fragment, though it too is capable of cleaving the vesicle or plasma-membran associated protein. It is conveniently obtained by insertion, deletion and/or substitution of a light chain or fragment thereof. In embodiments of the invention described below a variant sequence comprises (i) an N-terminal extension to a clostridial toxin light chain or fragment (ii) a clostridial toxin light chain or fragm nt modified by alteration of at least one amino acid (iii) a C-terminal extension to a clostridial toxin light chain or fragment, or (iv) combinations of 2 or more of (i)-(iii).

In further embodiments of the invention, the variant contains an amino a id sequence modified so that (a) there is no protease sensitive region between th LC and H_N components of the polypeptide, or (b) the protease sensitive region is specific for a particular protease. This latter embodiment is of use if it is desired to activate the endopeptidase activity of the light chain in a particular environment or cell. Though, in general, the polypeptides of the invention are activated prior to administration.

The first domain preferably exhibits endopeptidase activity specific for a substrat selected from one or more of SNAP-25, synaptobrevin/VAMP and syntaxin. Th cl stridial toxin is pref rably botulinum toxin or tetanus toxin.

In an embodim nt of the inv ntion d scribed in an xample bel w, th toxin light

chain and th portion of th t xin heavy chain are of botulinum t xin typ A. In a further embodiment of the invention described in an example below, the t xin light chain and the portion of the toxin heavy chain are of botulinum toxin typ B. The polypeptide optionally comprises a light chain or fragment or variant of on toxin type and a heavy chain or fragment or variant of another toxin type.

In a polypeptide according to the invention said second domain preferably comprises a clostridial toxin heavy chain H_N portion or a fragment or variant f a clostridial toxin heavy chain H_N portion. The fragment is optionally an N-terminal or C-terminal or internal fragment, so long as it retains the function of the H_N domain. Teachings of regions within the H_N responsible for its function ar provided for example in Biochemistry 1995, 34, pages 15175-15181 and Eur. J. Biochem, 1989, 185, pages 197-203. The variant has a different sequence fr m the H_N domain or fragment, though it too retains the function of the H_N domain. It is conveniently obtained by insertion, deletion and/or substitution of a H_N domain or fragment thereof. In embodiments of the invention, described below, it comprises (i) an N-terminal extension to a H_N domain or fragment, (ii) a C-terminal extension to a H_N domain or fragment, (iii) a modification to a H_N domain or fragment by alteration of at least one amino acid, or (iv) combinations of 2 or mor of (i)-(iii). The clostridial toxin is preferably botulinum toxin or tetanus toxin.

The invention also provides a polypeptide comprising a clostridial neurotoxin light chain and a N-terminal fragment of a clostridial neurotoxin heavy chain, th fragment preferably comprising at least 423 of the N-terminal amino acids of th heavy chain of botulinum toxin type A, 417 of the N-terminal amino acids of th heavy chain of botulinum toxin type B or the equivalent number of N-terminal amino acids of the heavy chain of other types of clostridial toxin such that the fragment possesses an equivalent alignment of homologous amino acid residues.

Thes p lyp ptid s of th invention ar thus n t comp s d of two or m r polyp ptides, link d f r example by di-sulphide bridg s into c mposit m lecules. Inst ad, these p lyp ptid s are singl chains and ar not activ r th ir activity is

significantly reduced in an *in vitro* assay of neurotoxin endop ptidase activity.

Further, the polypeptides may be susceptible to be converted into a form exhibiting endopeptidase activity by the action of a proteolytic agent, such as trypsin. In this way it is possible to control the endopeptidase activity of the toxin light chain.

In a specific embodiment of the invention described in an example below, there is provided a polypeptide lacking a portion designated H_C of a clostridial toxin heavy chain. This portion, seen in the naturally produced toxin, is responsible for binding of toxin to cell surface receptors prior to internalisation of the toxin. This specific embodiment is therefore adapted so that it can not be converted into active t xin, for example by the action of a proteolytic enzyme. The invention thus also provides a polypeptide comprising a clostridial toxin light chain and a fragment f a clostridial toxin heavy chain, said fragment being not capable of binding to th se cell surface receptors involved in the intoxicating action of clostridial toxin, and it is preferred that such a polypeptide lacks an intact portion designated H_C f a clostridial toxin heavy chain.

In further embodiments of the invention there are provided compositions containing a polypeptide comprising a clostridial toxin light chain and a portion designated H_N of a clostridial toxin heavy chain, and wherein the composition is free of clostridial toxin and free of any clostridial toxin precursor that may be converted into clostridial toxin by the action of a proteolytic enzyme. Examples of th se compositions include those containing toxin light chain and H_N sequences of botulinum toxin types A, B, C_1, D, E, F and G.

The polypeptides of the invention are conveniently adapted to bind to, or includ , a ligand for targeting to desired cells. The polypeptide optionally compris s a sequence that binds to, for example, an immunoglobulin. A suitable sequenc is a tandem rep at synth tic IgG binding domain d rived fr m d main B of Staphyloc ccal pr tein A. Choice f immun gl bulin specificity th n d termines the target for a polypeptid - immunogl bulin c mpl x. Alternativ ly, th

polypeptide c mpris s a non-clostridial sequence that binds to a cell surface receptor, suitable sequences including insulin-like growth factor-1 (IGF-1) which binds to its specific receptor on particular cell types and the 14 amino acid residue sequence from the carboxy-terminus of cholera toxin A subunit which is able to bind the cholera toxin B subunit and thence to GM1 gangliosides. A polypeptid according to the invention thus, optionally, further comprises a third domain adapted for binding of the polypeptide to a cell.

In a second aspect the invention provides a fusion protein comprising a fusion f (a) a polypeptide of the invention as described above with (b) a second polypeptide adapted for binding to a chromatography matrix so as to enable purification of th fusion protein using said chromatography matrix. It is convenient for the second polypeptide to be adapted to bind to an affinity matrix, such as a glutathione Sepharose, enabling rapid separation and purification of the fusion protein from an impure source, such as a cell extract or supernatant.

One possible second purification polypeptide is glutathione-S-transferase (GST), and others will be apparent to a person of skill in the art, being chosen so as to enable purification on a chromatography column according to conventional techniques.

As noted above, by proteolytic treatment, for example using trypsin, of a polypeptide of the invention it is possible to induce endopeptidase activity in th treated polypeptide. A third aspect of the invention provides a compositi n comprising a derivative of a clostridial toxin, said derivative retaining at least 10% of the endopeptidase activity of the clostridial toxin, said derivative further being non-toxic *in vivo* due to its inability to bind to cell surface receptors, and wher in the composition is free of any component, such as toxin or a further toxin derivative, that is toxic *in vivo*. The activity of the derivative preferably approach s that of natural t xin, and is thus pr f rably at least 30% and m st pr f rably at least 60% of natural toxin. Th ov rall endopeptidas activity f th c mp sition will, of c urs , als b determined by the am unt of th d rivative that is pr s nt.

While it is known to treat naturally produced clostridial toxin to remove the H_c domain, this treatment does not totally remove toxicity of the preparation, inst ad some residual toxin activity remains. Natural toxin treated in this way is therefor still not entirely safe. The composition of the invention, derived by treatment of a pure source of polypeptide advantageously is free of toxicity, and can conveniently be used as a positive control in a toxin assay, as a vaccine against clostridial toxin or for other purposes where it is essential that there is no residual toxicity in the composition.

The invention enables production of the polypeptides and fusion proteins of th invention by recombinant means.

A fourth aspect of the invention provides a nucleic acid encoding a polypeptide or a fusion protein according to any of the aspects of the invention described abov .

In one embodiment of this aspect of the invention, a DNA sequence provided t code for the polypeptide or fusion protein is not derived from native clostridial sequences, but is an artificially derived sequence not preexisting in nature.

A specific DNA (SEQ ID NO: 1) described in more detail below encod s a polypeptide or a fusion protein comprising nucleotides encoding residues 1-871 of a botulinum toxin type A. Said polypeptide comprises the light chain domain and the first 423 amino acid residues of the amino terminal portion of a botulinum t xin type A heavy chain. This recombinant product is designated LH_{423}/A (SEQ ID NO: 2).

In a second embodiment of this aspect of the invention a DNA sequence which codes for the polypeptide or fusion protein is derived from native clostridial sequences but codes for a polypeptide or fusion protein not found in natur .

A specific DNA (SEQ ID NO: 19) d scribed in mor d tail b low encodes a polyp ptid or a fusion prot in and compris s nucle tid s nc ding residues 1-

1171 of a botulinum toxin type B. Said polypeptide comprises the light chain domain and the first 728 amino acid residues of the amino terminal protein of a botulinum type B heavy chain. This recombinant product is designated LH_{728}/B (SEQ ID NO: 20).

The invention thus also provides a method of manufacture of a polypeptide comprising expressing in a host cell a DNA according to the third aspect of the invention. The host cell is suitably not able to cleave a polypeptide or fusi n protein of the invention so as to separate light and heavy toxin chains; for example, a non-clostridial host.

The invention further provides a method of manufacture of a polypeptide comprising expressing in a host cell a DNA encoding a fusion protein as described above, purifying the fusion protein by elution through a chromatography column adapted to retain the fusion protein, eluting through said chromatography column a ligand adapted to displace the fusion protein and recovering the fusion protein. Production of substantially pure fusion protein is thus made possible. Likewise, th fusion protein is readily cleaved to yield a polypeptide of the invention, again in substantially pure form, as the second polypeptide may conveniently be removed using the same type of chromatography column.

The LH_N/A derived from dichain native toxin requires extended digestion with trypsin to remove the C-terminal 1/2 of the heavy chain, the H_C domain. The loss of this domain effectively renders the toxin inactive *in vivo* by preventing its interaction with host target cells. There is, however, a residual toxic activity whi h may indicate a contaminating, trypsin insensitive, form of the whole type A neurotoxin.

In contrast, the recombinant preparations of the invention are the product of a discr et, defined g n c ding sequence and can n t be contaminated by full l ngth t xin pr t in. Furth rmore, th product as r cov red from *E. coli*, and fr m th r recombinant expressi n hosts, is an inactiv singl chain peptid r if xpr ssion

hosts produce a processed, active polyp ptide it is not a toxin. Endopeptidase activity of LH_{423}/A, as assessed by the current *in vitro* peptide cleavage assay, is wholly dependent on activation of the recombinant molecule between residues 430 and 454 by trypsin. Other proteolytic enzymes that cleave between these two residues are generally also suitable for activation of the recombinant mol cule. Trypsin cleaves the peptide bond C-terminal to Arginine or C-terminal to Lysine and is suitable as these residues are found in the 430-454 region and are exposed (s e Fig. 12).

The recombinant polypeptides of the invention are potential therapeutic agents for targeting to cells expressing the relevant substrate but which are not implicat d in effecting botulism. An example might be where secretion of neurotransmitter is inappropriate or undesirable or alternatively where a neuronal cell is hyperactive in terms of regulated secretion of substances other than neurotransmitter. In such an example the function of the H_C domain of the native toxin could be replaced by an alternative targeting sequence providing, for example, a cell receptor ligand and/ r translocation domain.

One application of the recombinant polypeptides of the invention will be as a reagent component for synthesis of therapeutic molecules, such as disclosed in WO-A-94/21300. The recombinant product will also find application as a non-t xic standard for the assessment and development of *in vitro* assays for detection of functional botulinum or tetanus neurotoxins either in foodstuffs or in environmental samples, for example as disclosed in EP-A-0763131.

A further option is addition, to the C-terminal end of a polypeptide of the inventi n, of a peptide sequence which allows specific chemical conjugation to targeting ligands of both protein and non-protein origin.

In y t a furth r mbodiment an alt rnativ targ ting ligand is add d t th N-terminus of polyp ptid s of th invention. R c mbinant LH_N derivativ s have be n d signat d that have specific proteas cleavag sites engine r d at th C-t rminus

of the LC at th putativ trypsin sensitiv region and also at th extreme C-terminus of the complete protein product. These sites will enhance the activational specificity of the recombinant product such that the dichain species can only b activated by proteolytic cleavage of a more predictable nature than use of trypsin.

The LH_N enzymatically produced from native BoNT/A is an efficient immunogen and thus the recombinant form with its total divorce from any full length neurot xin represents a vaccine component. The recombinant product may serve as a basal reagent for creating defined protein modifications in support of any of the abov areas.

Recombinant constructs are assigned distinguishing names on the basis of their amino acid sequence length and their Light Chain (L-chain, L) and Heavy Chain (H-chain, H) content as these relate to translated DNA sequences in the public domain or specifically to SEQ ID NO: 2 and SEQ ID NO: 20. The 'LH' designation is followed by '/X' where 'X' denotes the corresponding clostridial toxin serotype or class, e.g. 'A' for botulinum neurotoxin type A or 'TeTx' for tetanus toxin. Sequence variants from that of the native toxin polypeptide are given in parenth sis in standard format, namely the residue position number prefixed by the residue of the native sequence and suffixed by the residue of the variant.

Subscript number prefixes indicate an amino-terminal (N-terminal) extension, or where negative a deletion, to the translated sequence. Similarly, subscript numb r suffixes indicate a carboxy terminal (C-terminal) extension or where negativ numbers are used, a deletion. Specific sequence inserts such as protease cleavage sites are indicated using abbreviations, e.g. Factor Xa is abbreviated to FXa. L-chain C-terminal suffixes and H-chain N-terminal prefixes are separated by a '/' to indicate the predicted junction between the L and H-chains. Abbreviations for engineered ligand sequences are prefixed or suffixed to the clostridial L-chain or H-chain corresponding to their position in th translati n product.

Following this nom nclatur ,

LH_{423}/A = SEQ ID NO: 2, containing the ntire L-chain and 423 amino acids of the H-chain of botulinum neurotoxin typ A;

$_2LH_{423}/A$ = a variant of this molecule, containing a two amino acid extension to the N-terminus of the L-chain;

$_2L_{/2}H_{423}/A$ = a further variant in which the molecule contains a two amino acid extension on the N-terminus of both th L-chain and the H-chain;

$_2L_{FXa/2}H_{423}/A$ = a further variant containing a two amino acid extensi n to the N-terminus of the L-chain, and a Factor Xa cleavage sequence at the C-terminus of the L-chain which, after cleavage of the molecule with Fact r Xa leaves a two amino acid N-terminal extension to th H-chain component; and

$_2L_{FXa/2}H_{423}/A$-IGF-1 = a variant of this molecule which has a further C-t rminal extension to the H-chain, in this example the insulin-like growth factor 1 (IGF-1) sequence.

There now follows description of specific embodiments of the invention, illustrated by drawings in which:

Fig. 1 shows a schematic representation of the domain structure of botulinum neurotoxin type A (BoNT/A);

Fig. 2 sh ws a sch matic r pr sentation f assembly f th g ne for an mb diment of the inv nti n d signat d LH_{423}/A;

Fig. 3 is a graph comparing activity of native toxin, trypsin generated "native" LH_N/A and an embodiment of the invention designated $_2LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$) in an *in vitro* peptide cleavage assay;

Fig. 4 is a comparison of the first 33 amino acids in publish d sequences of native toxin and embodiments of the invention;

Fig. 5 shows the transition region of an embodiment of th invention designated $L/_4H_{423}/A$ illustrating insertion of four amino acids at the N-terminus of the H_N sequence; amino acids coded for by the *Eco* 47 III restricti n endonuclease cleavage site are marked and the H_N sequence then begins ALN...;

Fig. 6 shows the transition region of an embodiment of the invention designated $L_{FXa/3}H_{423}/A$ illustrating insertion of a Factor Xa cleavage site at the C-terminus of the L-chain, and three additional amino acids coded for at the N-terminus of the H-sequence; the N-terminal amino acid of the cleavage-activat d H_N will be cysteine;

Fig. 7 shows the C-terminal portion of the amino acid sequence of an embodiment of the invention designated $L_{FXa/3}H_{423}/A$-IGF-1, a fusion protein; the IGF-1 sequence begins at position G_{882};

Fig. 8 shows the C-terminal portion of the amino acid sequence of an embodiment of the invention designated $L_{FXa/3}H_{423}/A$-CtxA14, a fusion protein; the C-terminal CtxA sequence begins at p sition Q_{882};

Fig.9 shows the C-terminal portion f the amin acid sequ nce ofan

embodiment of the invention designated $L_{FXa/3}H_{423}$/A-ZZ, a fusion protein; the C-terminal ZZ sequence begins at position A_{890} immediately after a genenase recognition site (underlin d);

show schematic representations of manipulations of

Figs. 10 & 11 polypeptides of the invention; Fig. 10 shows LH_{423}/A with N-terminal addition of an affinity purification peptide (in this case GST) and C-terminal addition of an Ig binding domain; protease cleavage sites R1, R2 and R3 enable selective enzymatic separation of domains; Fig. 11 shows specific examples of protease cleavag sites R1, R2 and R3 and a C-terminal fusion p ptide sequence;

Fig. 12 shows the trypsin sensitive activation region f a polypeptide of the invention;

Fig. 13 shows Western blot analysis of recombinant LH_{107}/B expressed from *E.coli*; panel A was probed with anti-BoNT/B antiserum; Lane 1, molecular weight standards; lanes 2 & 3, native BoNT/B; lane 4, immunopurifi d LH_{107}/B; panel B was probed with anti-T7 peptid tag antiserum; lane 1, molecular weight standards; lan s 2 & 3, positive control *E.coli* T7 expression; lan 4 immunopurified LH_{107}/B.

The sequence listing that accompanies this application contains the following sequences:-

SEQ ID NO:	Sequence
1	DNA coding for LH_{23}/A

2	LH_{423}/A
3	DNA coding for $_{23}LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$), of which an N-terminal portion is shown in Fig. 4.
4	$_{23}LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$)
5	DNA coding for $_2LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$), of which an N-terminal portion is shown in Fig.4
6	$_2LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$)
7	DNA coding for native BoNT/A according to Binz et al
8	native BoNT/A according to Binz et al
9	DNA coding for $L_{/4}H_{423}/A$
10	$L_{/4}H_{423}/A$
11	DNA coding for $L_{FXa}/_3H_{423}/A$
12	$L_{FXa}/_3H_{423}/A$
13	DNA coding for $L_{FXa}/_3H_{423}/A$-IGF-1
14	$L_{FXa}/_3H_{423}/A$-IGF-1
15	DNA coding for $L_{FXa}/_3H_{423}/A$-CtxA14
16	$L_{FXa}/_3H_{423}/A$-CtxA14
17	DNA coding for $L_{FXa/3}H_{423}/A$-ZZ
18	$L_{FXa/3}H_{423}/A$-ZZ
19	DNA coding for LH_{728}/B
20	LH_{728}/B
21	DNA coding for LH_{417}/B
22	LH_{417}/B
23	DNA coding for LH_{107}/B
24	LH_{107}/B
25	DNA coding for LH_{423}/A ($Q_2E,N_{26}K,A_{27}Y$)
26	LH_{423}/A ($Q_2E,N_{26}K,A_{27}Y$)
27	DNA coding f r LH_{417}/B wher in th first 274 bases are

modifi d t hav an *E.coli* c don bias

28 DNA coding for LH_{417}/B wherein bases 691-1641 of th native BoNT/B sequence have been replaced by a degenerate DNA coding for amino acid residues 231-547 of the native BoNT/B polypeptide

Example 1

A 2616 base pair, double stranded gene sequence (SEQ ID NO: 1) has been assembled from a combination of synthetic, chromosomal and polymerase-chain-reaction generated DNA (Figure 2). The gene codes f r a polypeptide of 871 amino acid residues corresponding to the entire light-chain (LC, 448 amino acids) and 423 residues of the amino terminus of the heavy-chain (H_C) of botulinum neurotoxin type A. This recombinant product is designated the LH_{423}/A fragment (SEQ ID NO: 2).

Construction of the recombinant product

The first 918 base pairs of the recombinant gene were synthesised by concatenation of short oligonucleotides to generate a coding sequence with an *E. coli* codon bias. Both DNA strands in this region were completely synthesis d as short overlapping oligonucleotides which were phosphorylated, annealed and ligated to generate the full synthetic region ending with a unique *KpnI* restricti n site. The remainder of the LH_{423}/A coding sequence was PCR amplified from total chromosomal DNA from *Clostridium botulinum* and annealed to the synth tic portion of the gene.

The internal PCR amplified product sequences were then deleted and replaced with the native, fully sequenced, regions from clones of *C. botulinum* chromosomal origin t g nerat th final gen c nstru t. Th final comp siti n is synth tic DNA (bases 1-913), p lym rase amplifi d DNA (bas s 914-1138 and 1976-2616) and th remainder is f *C. botulinum* chrom somal origin (bases 1139-1975). Th

assembled gen was then fully sequenced and cloned into a variety of *E.coli* plasmid vectors for expression analysis.

Expression of the recombinant gene and recovery of protein product

The DNA is expressed in *E. coli* as a single nucleic acid transcript producing a soluble single chain polypeptide of 99,951 Daltons predicted molecular weight. Th gene is currently expressed in *E. coli* as a fusion to the commercially availabl coding sequence of glutathione S-transferase (GST) of *Schistosoma japonicum* but any of an extensive range of recombinant gene expression vectors such as pEZZ18, pTrc99, pFLAG or the pMAL series may be equally effective as might expression in other prokaryotic or eukaryotic hosts such as the Gram positive bacilli, the yeast *P. pastoris* or in insect or mammalian cells under appropriate conditions.

Currently, *E. coli* harbouring the expression construct is grown in Luria-Bertani broth (L-broth pH 7.0, containing 10 g/l bacto-tryptone, 5 g/l bacto-yeast extract and 10 g/l sodium chloride) at 37° C until the cell density (biomass) has an optical absorbance of 0.4- 0.6 at 600 nm and the cells are in mid-logarithmic growth phase. Expression of the gene is then induced by addition of isopropylthio-β-D-galactosidase (IPTG) to a final concentration of 0.5 mM. Recombinant gene expression is allowed to proceed for 90 min at a reduced temperature of 25°C. The cells are then harvested by centrifugation, ar resuspended in a buffer solution containing 10 mM Na_2HPO_4, 0.5 M NaCl, 10 mM EGTA, 0.25% Tween, pH 7.0 and then frozen at -20°C. For extraction of th recombinant protein the cells are disrupted by sonication. The cell extract is then cleared of debris by centrifugation and the cleared supernatant fluid containing soluble recombinant fusion protein (GST- LH_{423}/A) is stored at -20°C pending purification. A proportion of recombinant material is not released by the sonication procedure and this probably reflects insolubility or inclusion b dy f rmati n. Curr ntly we d n t extract this material f r analysis but if desir d this c uld be readily achi ved using m thods known t th se skill d in th art.

The recombinant GST- LH_{423}/A is purified by adsorption onto a commercially prepared affinity matrix of glutathione Sepharose and subsequent elution with reduced glutathione. The GST affinity purification marker is then removed by proteolytic cleavage and reabsorption to glutathione Sepharose; recombinant LH_{423}/A is recovered in the non-adsorbed material.

Construct variants

A variant of the molecule, LH_{423}/A ($Q_2E,N_{26}K,A_{27}Y$) (SEQ ID NO: 26) has b en produced in which three amino acid residues have been modified within the light chain of LH_{423}/A producing a polypeptide containing a light chain sequence different to that of the published amino acid sequence of the light chain of BoNT/A .

Two further variants of the gene sequence that have been expressed and th corresponding products purified are $_{23}LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$) (SEQ ID NO: 4) which has a 23 amino acid N-terminal extension as compared to the predicted native L-chain of BoNT/A and $_{2}LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$) (SEQ ID NO: 6) which has a 2 amino acid N-terminal extension (Figure 4).

In yet another variant a gene has been produced which contains a *Eco* 47 III restriction site between nucleotides 1344 and 1345 of the gene sequence given in (SEQ ID NO: 1). This modification provides a restriction site at the position in th gene representing the interface of the heavy and light chains in native neurotoxin, and provides the capability to make insertions at this point using standard restriction enzyme methodologies known to those skilled in the art. It will als b obvious to those skilled in the art that any one of a number of restriction sites could be so employed, and that the *Eco* 47 III insertion simply exemplifies this approach. Similarly, it would be obvious for one skilled in the art that insertion of a restriction site in the manner described could be performed on any gene of the inventi n. The g n d scribed, when expressed, c d s for a polyp ptid , $L_{/4}H_{423}/A$ (SEQ ID NO: 10), which c ntains an additional f ur amin acids b tw n amino acids 448 and 449 of LH $_{23}$/A at a p sition quivalent to the amin terminus f the

heavy chain of native BoNT/A.

A variant of the gene has been expressed, $L_{FXa/3}H_{423}/A$ (SEQ ID NO: 12), in which a specific proteolytic cleavage site was incorporated at the carboxy-terminal end of the light chain domain, specifically after residue 448 of $L_{/4}H_{423}/A$. The cleavage site incorporated was for Factor Xa protease and was coded for by modification of SEQ ID NO: 1. It will be apparent to one skilled in the art that a cleavage sit for another specified protease could be similarly incorporated, and that any gene sequence coding for the required cleavage site could be employed. Modification of the gene sequence in this manner to code for a defined protease site could b performed on any gene of the invention.

Variants of $L_{FXa/3}H_{423}/A$ have been constructed in which a third domain is present at the carboxy-terminal end of the polypeptide which incorporates a specific binding activity into the polypeptide.

Specific examples described are:

(1) $L_{FXa/3}H_{423}/A$-IGF-1 (SEQ ID NO: 14), in which the carboxy-terminal domain has a sequence equivalent to that of insulin-like growth factor-1 (IGF-1) and is able to bind to the insulin-like growth factor receptor with high affinity;

(2) $L_{FXa/3}H_{423}/A$-CtxA14 (SEQ ID NO: 16), in which the carboxy-terminal domain has a sequence equivalent to that of the 14 amino acids from the carboxy-terminus of the A-subunit of cholera toxin (CtxA) and is thereby able to interact with the cholera toxin B-subunit pentamer; and

(3) $L_{FXa/3}H_{423}/A$-ZZ (SEQ ID NO: 18), in which the carboxy-terminal domain is a tandem repeating synthetic IgG binding domain. This variant also exemplifies another modification applicabl t the current inventi n, nam ly th inclusi n in the g n of a s quenc coding f r a pr teas cl avag sit locat d betw en th end of th clostridial h avy chain s qu nc and th s qu nc c ding for th binding

ligand. Specifically in this xample a s quence is inserted at nucleotides 2650 to 2666 coding for a genenase cleavage site. Expression of this gene produces a polypeptide which has the desired protease sensitivity at the interface between the domain providing H_N function and the binding domain. Such a modification enables selective removal of the C-terminal binding domain by treatment of the polypeptide with the relevant protease.

It will be apparent that any one of a number of such binding domains could be incorporated into the polypeptide sequences of this invention and that the ab v examples are merely to exemplify the concept. Similarly, such binding domains can be incorporated into any of the polypeptide sequences that are the basis of this invention. Further, it should be noted that such binding domains c uld be incorporated at any appropriate location within the polypeptide molecules f the invention.

Further embodiments of the invention are thus illustrated by a DNA of the invention further comprising a desired restriction endonuclease site at a desired location and by a polypeptide of the invention further comprising a desired protease cl avag site at a desired location.

The restriction endonuclease site may be introduced so as to facilitate further manipulation of the DNA in manufacture of an expression vector for expressing a polypeptide of the invention; it may be introduced as a consequence of a pr vi us step in manufacture of the DNA; it may be introduced by way of modificati n by insertion, substitution or deletion of a known sequence. The consequenc of modification of the DNA may be that the amino acid sequence is unchang d, or may be that the amino acid sequence is changed, for example resulting in introduction of a desired protease cleavage site, either way the polypeptide retains its first and second domains having the properties required by the inventi n.

Figur 10 is a diagrammatic repr sentation f an expression product ex mplifying featur s d scrib d in this example. Sp cifically, it illustrat s a singl polyp ptid

incorporating a domain quivalent to th light chain of botulinum n urotoxin type A and a domain equivalent to the H_N domain of the heavy chain of botulinum neurotoxin type A with a N-terminal extension providing an affinity purification domain, namely GST, and a C-terminal extension providing a ligand binding domain, namely an IgG binding domain. The domains of the polypeptide are spatially separated by specific protease cleavage sites enabling selective enzymatic separation of domains as exemplified in the Figure. This concept is more specifically depicted in Figure 11 where the various protease sensitivities are defined for the purpose of example.

Assay of product activity

The LC of botulinum neurotoxin type A exerts a zinc-dependent endopeptidas activity on the synaptic vesicle associated protein SNAP-25 which it cleaves in a specific manner at a single peptide bond. The $_2LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$) (SEQ ID NO: 6) cleaves a synthetic SNAP-25 substrate *in vitro* under the same conditions as the native toxin (Figure 3). Thus, the modification of the polypeptide sequ nc of $_2LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$) relative to the native sequence and within the minimal functional LC domains does not prevent the functional activity of the LC domains.

This activity is dependent on proteolytic modification of the recombinant GST-$_2LH_{423}/A$ ($Q_2E,N_{26}K,A_{27}Y$) to convert the single chain polypeptide product t a disulphide linked dichain species. This is currently done using the proteolytic enzyme trypsin. The recombinant product (100-600 μg/ml) is incubated at 37°C for 10-50 minutes with trypsin (10 μg/ml) in a solution containing 140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3. The reaction is terminated by addition of a 100-fold molar excess of trypsin inhibitor. Th activation by trypsin generates a disulphide linked dichain species as determined by polyacrylamide gel electrophoresis and immunoblotting analysis using polycl nal anti-botulinum neur t xin type A antiserum.

$_2LH_{423}/A$ is more stable in the pr ence of trypsin and m r active in the *in vitro*

peptid cleavag assay than is $_{23}LH_{423}/A$. Both variants, however, are fully functional in the *in vitro* peptide cleavage assay. This demonstrates that the recombinant molecule will tolerate N-terminal amino acid extensions and this may be expanded to other chemical or organic moieties as would be obvious to thos skilled in the art.

Example 2

As a further exemplification of this invention a number of gene sequences have been assembled coding for polypeptides corresponding to the entire light-chain and varying numbers of residues from the amino terminal end of the heavy chain of botulinum neurotoxin type B. In this exemplification of the disclosure th gen sequences assembled were obtained from a combination of chromosomal and polymerase-chain-reaction generated DNA, and therefore have the nucle tid sequence of the equivalent regions of the natural genes, thus exemplifying th principle that the substance of this disclosure can be based upon natural as well as a synthetic gene sequences.

The gene sequences relating to this example were all assembled and expr ss d using methodologies as detailed in Sambrook J, Fritsch E F & Maniatis T (1989) Molecular Cloning: A Laboratory Manual (2nd Edition), Ford N, Nolan C, Ferguson M & Ockler M (eds), Cold Spring Harbor Laboratory Press, New York, and known to those skilled in the art.

A gene has been assembled coding for a polypeptide of 1171 amino acids corresponding to the entire light-chain (443 amino acids) and 728 residues from th amino terminus of the heavy chain of neurotoxin type B. Expression of this g n produces a polypeptide, LH_{728}/B (SEQ ID NO: 20), which lacks the specific neuronal binding activity of full length BoNT/B.

A gen has als b en ass mbl d coding f r a variant polypeptid , LH_{417}/B (SEQ ID NO: 22), which poss ss s an amino acid s qu nc at its carb xy terminus

equivalent by amino acid homology to that at the carboxy-terminus of th heavy chain fragment in native LH_N/A .

A gene has also been assembled coding for a variant polypeptide, LH_{107}/B (SEQ ID NO: 24) , which expresses at its carboxy-terminus a short sequence from th amino terminus of the heavy chain of BoNT/B sufficient to maintain solubility of the expressed polypeptide.

Construct Variants

A variant of the coding sequence for the first 274 bases of the gene shown in SEQ ID NO: 21 has been produced which whilst being a non-native nucleotide sequence still codes for the native polypeptide.

Two double stranded, a 268 base pair and a 951 base pair, gene sequences hav been created using an overlapping primer PCR strategy. The nucleotide bias of these sequences was designed to have an *E.coli* codon usage bias.

For the first sequence, six oligonucleotides representing the first (5') 268 nucleotides of the native sequence for botulinum toxin type B were synthesis d. For the second sequence 23 oligonucleotides representing internal sequ nce nucleotides 691-1641 of the native sequence for botulinum toxin type B were synthesised. The oligonucleotides ranged from 57-73 nucleotides in length. Overlapping regions, 17-20 nucleotides, were designed to give melting temperatures in the range 52-56°C. In addition, terminal restriction endonuclease sites of the synthetic products were constructed to facilitate insertion of these products into the exact corresponding region of the native sequence. The 268 bp 5' synthetic sequence has been incorporated into the gene shown in SEQ ID NO: 21 in place of the original first 268 bases (and is shown in SEQ ID NO: 27). Similarly the s quenc could b insert d int oth r genes of th xampl s.

Anoth r variant s qu nc equival nt t nucl otides 691 to 1641 of SEQ ID NO: 21

, and employing non-native codon usage whilst c ding for a native polypeptide sequence, has been constructed using the internal synthetic sequence. This sequence (SEQ ID NO: 28) can be incorporated, alone or in combination with other variant sequences, in place of the equivalent coding sequence in any of th gen s of the example.

Example 3

An exemplification of the utility of this invention is as a non-toxic and effectiv immunogen. The non-toxic nature of the recombinant, single chain material was demonstrated by intraperitoneal administration in mice of GST-$_2$LH$_{423}$/A. The polypeptide was prepared and purified as described above. The amount f immunoreactive material in the final preparation was determined by enzym linked immunosorbent assay (ELISA) using a monoclonal antibody (BA11) reactive against a conformation dependent epitope on the native LH$_N$/A. The recombinant mat rial was serially diluted in phosphate buffered saline (PBS; NaCl 8 g/l, KCl 0.2 g/l, Na_2HPO_4 1.15 g/l, KH_2PO_4 0.2 g/l, pH 7.4) and 0.5 ml volumes injected into 3 groups of 4 mice such that each group of mice received 10, 5 and 1 micrograms of material respectively. Mice were observed for 4 days and no deaths were se n.

For immunisation, 20 μg of GST-$_2$LH$_{423}$/A in a 1.0 ml volume of water-in-oil emulsion (1:1 vol:vol) using Freund's complete (primary injections only) or Fr und's incomplete adjuvant was administered into guinea pigs via two sub-cutan ou dorsal injections. Three injections at 10 day intervals were given (day 1, day 10 and day 20) and antiserum collected on day 30. The antisera were shown by ELISA to be immunoreactive against native botulinum neurotoxin type A and to its derivative LH$_N$/A. Antisera which were botulinum neurotoxin reactive at a dilution of 1:2000 were used for evaluation of neutralising efficacy in mice. For neutralisation assays 0.1 ml of antiserum was diluted into 2.5 ml of gelatin phosphat buff r (GPB; Na_2HPO_4 anhydr us 10 g/l, g latin (Difc) 2 g/l, pH 6.5-6.6) containing a diluti n rang from 0.5 μg ($5X10^{-6}$ g) to 5 picograms ($5X10^{-12}$ g). Aliquots of 0.5 ml w r injected into mice intraperit n ally and deaths r cord d

over a 4 day period. Th results ar shown in Table 1 and Tabl 2. It can clearly be seen that 0.5 ml of 1:40 diluted anti- GST-$_2LH_{423}$/A antiserum can protect mice against intraperitoneal challenge with botulinum neurotoxin in the range 5 pg - 50 ng (1 - 10,000 mouse LD50; 1 mouse LD50 = 5 pg).

TABLE 1. Neutralisation of botulinum neurotoxin in mice by guinea pig anti-GST-$_2$LH$_{423}$/A antiserum.

Survivors On Day	Botulinum Toxin/mouse 0.5μg	0.005μg	0.0005μg	0.5ng	0.005ng	5pg	Control (no toxin)
1	0	4	4	4	4	4	4
2	-	4	4	4	4	4	4
3	-	4	4	4	4	4	4
4	-	4	4	4	4	4	4

TABLE 2. Neutralisation of botulinum neurotoxin in mice by non-immune guin a pig antiserum.

Survivors On Day	Botulinum Toxin/mouse 0.5μg	0.005μg	0.0005μg	0.5ng	0.005ng	5pg	Control (no toxin)
1	0	0	0	0	0	2	4
2	-	-	-	-	-	0	4
3	-	-	-	-	-	-	4
4	-	-	-	-	-	-	4

Example 4

Expression of recombinant LH$_{107}$/B in *E. coli*.

As an x mplificati n of th xpr ssi n of a nucleic acid coding f r a LH$_N$ f a cl stridial n urot xin of a s r typ th r than b tulinum neur t xin typ A, th nucl ic acid s qu nc (SEQ ID NO: 23) c ding for th polyp ptid LH$_{107}$/B (SEQ ID

NO: 24) was inserted int the commercially available plasmid pET28a (Novogen, Madison, WI, USA). The nucleic acid was expressed in *E. coli* BL21 (DE3) (New England BioLabs, Beverley, MA, USA) as a fusion protein with a N-terminal T7 fusion peptide, under IPTG induction at 1 mM for 90 minutes at 37°C. Cultur s were harvested and recombinant protein extracted as described previously for LH_{423}/A.

Recombinant protein was recovered and purified from bacterial paste lysates by immunoaffinity adsorption to an immobilised anti-T7 peptide monoclonal antibody using a T7 tag purification kit (New England bioLabs, Beverley, MA, USA). Purified recombinant protein was analysed by gradient (4-20%) denaturing SDS-polyacrylamide gel electrophoresis (Novex, San Diego, CA, USA) and west rn blotting using polyclonal anti-botulinum neurotoxin type antiserum or anti-T7 antiserum. Western blotting reagents were from Novex, immunostained proteins were visualised using the Enhanced Chemi-Luminescence system (ECL) from Amersham. The expression of an anti-T7 antibody and anti-botulinum neurotoxin type B antiserum reactive recombinant product is demonstrated in Figure 13.

The recombinant product was soluble and retained that part of the light chain responsible for endopeptidase activity.

The invention thus provides recombinant polypeptides useful inter alia as immunogens, enzyme standards and components for synthesis of molecules as described in WO-A-94/21300.

CLAIMS

1. A polypeptide comprising first and second domains, wherein said first domain is adapted to cleave one or more vesicle or plasma-membrane associat d proteins essential to exocytosis, and wherein said second domain is adapted (i) to translocate the polypeptide into a cell or (ii) to increase the solubility of th polypeptide compared to the solubility of the first domain on its own or (iii) both to translocate the polypeptide into a cell and to increase the solubility of th polypeptide compared to the solubility of the first domain on its own, said polypeptide being free of clostridial neurotoxin and free of clostridial neurot xin precursor that can be converted into toxin by proteolytic action.

2. A polypeptide according to Claim 1 wherein said first domain compris s a clostridial toxin light chain.

3. A polypeptide according to Claim 1 wherein said first domain comprises a fragment or variant of a clostridial toxin light chain.

4. A polypeptide according to Claim 2 or 3 wherein the clostridial toxin is a botulinum toxin.

5. A polypeptide according to any preceding claim wherein the first domain exhibits endopeptidase activity specific for a substrate selected from one or more of SNAP-25, synaptobrevin/VAMP and syntaxin.

6. A polypeptide according to any preceding claim wherein said second domain comprises a clostridial toxin heavy chain H_N portion.

7. A p lypeptid acc rding t any of Claim 1-5 wher in said s cond domain compris s a fragm nt or variant f a clostridial t xin h avy chain H_N portion.

8. A polyp ptid ac ording to Claim 6 or 7 wh r in th clostridial toxin is a

botulinum toxin.

9. A polypeptide according to any of Claims 1-8 further comprising a third domain adapted for binding of the polypeptide to a cell, by binding of the third domain directly to a cell or by binding of the third domain to a ligand or to ligands that bind to a cell.

10. A polypeptide according to Claim 9 wherein said third domain is for binding the polypeptide to an immunoglobulin.

11. A polypeptide according to Claim 10 wherein said third domain is a tandem repeat synthetic IgG binding domain derived from domain β of Staphylococcal protein A.

12. A polypeptide according to Claim 9 wherein said third domain comprises an amino acid sequence that binds to a cell surface receptor.

13. A polypeptide according to Claim 12 wherein said third domain is insulin-lik growth factor-1 (IGF-1).

14. A polypeptide according to any preceding claim comprising a botulinum toxin light chain or a fragment or a variant of a botulinum toxin light chain and a portion designated H_N of a botulinum toxin heavy chain.

15. A polypeptide according to Claim 14 wherein one or both of (a) the toxin light chain or fragment or variant of toxin light chain and (b) the portion of the toxin heavy chain are of botulinum toxin type A.

16. A polyp ptid acc rding to Claim 15 wh r in the botulinum t xin typ A light chain variant ha at r sidu 2 a glutamat , at r sidue 26 a lysin and at r sidu 27 a tyrosin .

17. A polypeptid according to Claim 14 wherein one or both of (a) th toxin light chain or fragment or variant of toxin light chain and (b) the portion of th toxin heavy chain are of botulinum toxin type B.

18. A polypeptide according to any of Claims 1-13 comprising a botulinum toxin light chain or a fragment or a variant of a botulinum toxin light chain and at least 100 N-terminal amino acids of a botulinum toxin heavy chain.

19. A polypeptide according to Claim 18 comprising a botulinum toxin typ B light chain, or a fragment or variant thereof, and 107 N-terminal amino acids of a botulinum toxin type B heavy chain.

20. A polypeptide according to Claim 15 or 16 comprising at least 423 of th N-terminal amino acids of botulinum toxin type A heavy chain.

21. A polypeptide according to Claim 20 comprising a botulinum toxin typ A light chain and 423 N-terminal amino acids of a botulinum toxin type A heavy chain.

22. A polypeptide according to Claim 20 comprising a botulinum toxin typ A light chain variant wherein residue 2 is a glutamate, residue 26 is a lysin and residue 27 is a tyrosine, and 423 N-terminal amino acids of a botulinum toxin typ A heavy chain.

23. A polypeptide according to Claim 17 comprising at least 417 of th N-terminal amino acids of botulinum toxin type B heavy chain.

24. A polypeptide according to Claim 23 comprising a botulinum toxin type B light chain and 417 N-t rminal amin acids of a b tulinum t xin typ B h avy chain.

25. A p lypeptide acc rding to any of Claims 14-24 lacking a p rtion d signat d

H_C of a botulinum toxin heavy chain.

26. A polypeptide comprising a botulinum toxin light chain and a fragment of a botulinum toxin heavy chain, said fragment being not capable of binding to cell surface receptors.

27. A polypeptide according to Claim 26 lacking an intact portion designated H_C of a botulinum toxin heavy chain.

28. A polypeptide according to any preceding claim comprising a variant of a clostridial toxin and further comprising a site for cleavage by a proteolytic enzym , which cleavage site is not present in the native toxin.

29. A polypeptide according to Claim 28 comprising a variant of a clostridial toxin light chain and further comprising a site for cleavage by a proteolytic enzyme, which cleavage site is not present in the native toxin light chain.

30. A polypeptide according to Claim 28 or 29 comprising a variant of a clostridial toxin heavy chain H_N portion and further comprising a site for cleavag by a proteolytic enzyme, which cleavage site is not present in the native toxin heavy chain H_N portion.

31. A polypeptide according to Claim 28, 29 or 30 obtainable by modificati n of a DNA encoding the polypeptide so as to introduce one or more nucleotid s coding for the cleavage site.

32. A fusion protein comprising a fusion of (a) a polypeptide according to any of Claims 1-31 with (b) a second polypeptide being a polypeptide or oligopeptide adapted for binding to an affinity matrix so as to nabl purificati n f th fusion prot in using said matrix.

33. A fusi n pr t in according t Claim 32 wh r in said sec nd p lypeptide is

adapted to bind to a chromatography column, such as an affinity matrix of glutathione Sepharose.

34. A fusion protein according to Claim 32 or 33 wherein a specific protease cleavage site is incorporated between the first and second polypeptides, said protease site enabling proteolytic separation of first and second polypeptid s.

35. A composition comprising a derivative of a clostridial toxin, said derivativ retaining at least 10% of the endopeptidase activity of the botulinum toxin, said derivative further being non-toxic *in vivo* due to its inability to bind to cell surfac receptors, and wherein the composition is free of any component, such as toxin. r a further toxin derivative, that is toxic *in vivo*.

36. A composition according to Claim 35 or a polypeptide according to any f Claims 1-31 or a fusion protein according to Claim 32, 33 or 34 for use as a positive control in a toxin assay.

37. A composition according to Claim 35 or a polypeptide according to any of Claims 1-31 or a fusion protein according to Claim 32, 33 or 34 for us as a vaccine against clostridial toxin.

38. A composition according to Claim 35 or a polypeptide according to any of Claims 1-31 or a fusion protein according to Claim 32, 33 or 34 for *in vivo* us .

39. A pharmaceutical composition comprising a composition according to Claim 35, a polypeptide according to any of claims 1-31 or a fusion protein according to Claim 32, 33 or 34, in combination with a pharmaceutically acceptable carrier.

40. A nucl ic acid encoding a p lyp ptid or a fusion protein according to any of Claims 1-34.

41. A nucl ic acid ncoding a p lypeptid r a fusi n prot in according t Claim

40 and comprising nucleotides encoding residues 1-448 of a botulinum toxin typ A light chain.

42. A nucleic acid according to Claim 40 or 41 comprising nucleotides encoding residues 1-423 of a botulinum toxin type A heavy chain H_N domain.

43. A nucleic acid encoding a polypeptide or a fusion protein according to Claim 40 and comprising nucleotides encoding residues 1-470 of a botulinum toxin type B light chain.

44. A nucleic acid encoding a polypeptide or a fusion protein according to Claim 40 or 43 comprising nucleotides encoding residues 1-417 of a botulinum toxin type B heavy chain H_N domain.

45. A nucleic acid according to any of Claims 40-44 comprising nucleotid s encoding a restriction endonuclease cleavage site not present in native clostridial toxin sequence.

46. A nucleotide according to Claim 45 obtainable by modification of a nucleotide encoding a polypeptide or fusion protein according to any of claims 1-34 so as to introduce said cleavage site.

47. A DNA according to any of claims 40-46.

48. A DNA selected from SEQ ID No:s 1, 8, 10, 12, 14, 16, 18, 23 and 24.

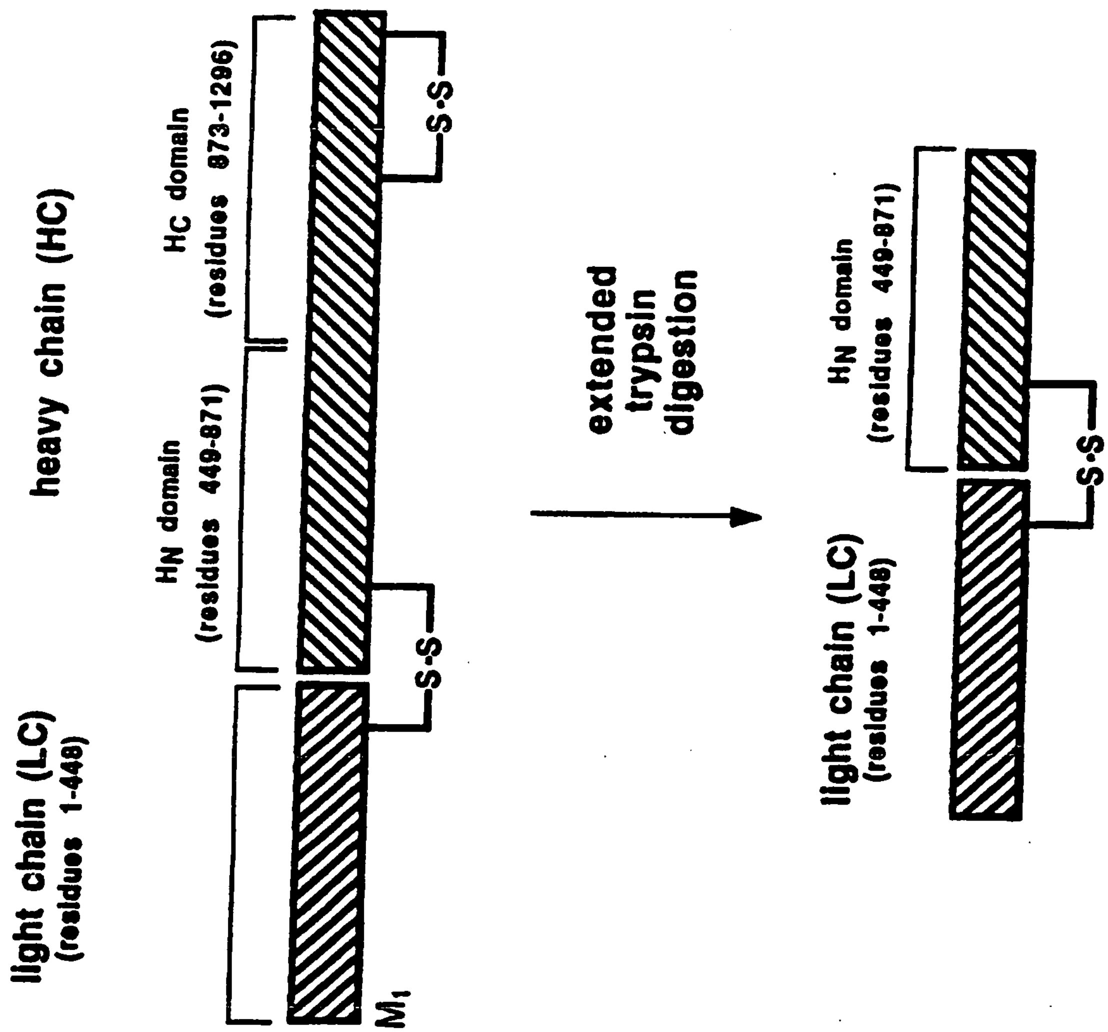
49. A method of manufacture of a polypeptide according to any of Claims 1-31 comprising expressing in a host cell a nucleic acid according to any of Claims 40-48 and r c v ring the p lyp ptid .

50. A m th d of manufactur of a polyp ptid acc rding to any of Claims 1-31 comprising xpr ssing in a h st c ll a nucl ic acid nc ding a fusion pr t in

according to Claim 32, 33 or 34, purifying the fusion protein by eluting th fusion protein through an affinity matrix adapted to retain the fusion protein and eluting through said matrix a ligand adapted to displace the fusion protein, and rec vering the fusion protein.

51. A method of manufacture according to Claims 49 or 50 in which the nucleic acid is DNA.

52. A cell expressing a polypeptide or fusion protein according to any of Claims 1-34.



heavy chain (HC)
light chain (LC)
(residues 1-448)
H_N domain
(residues 449-871)
H_C domain
(residues 873-1296)
S-S
S-S
M_1
extended
trypsin
digestion
light chain (LC)
(residues 1-448)
H_N domain
(residues 449-871)
S-S

FIG. 1

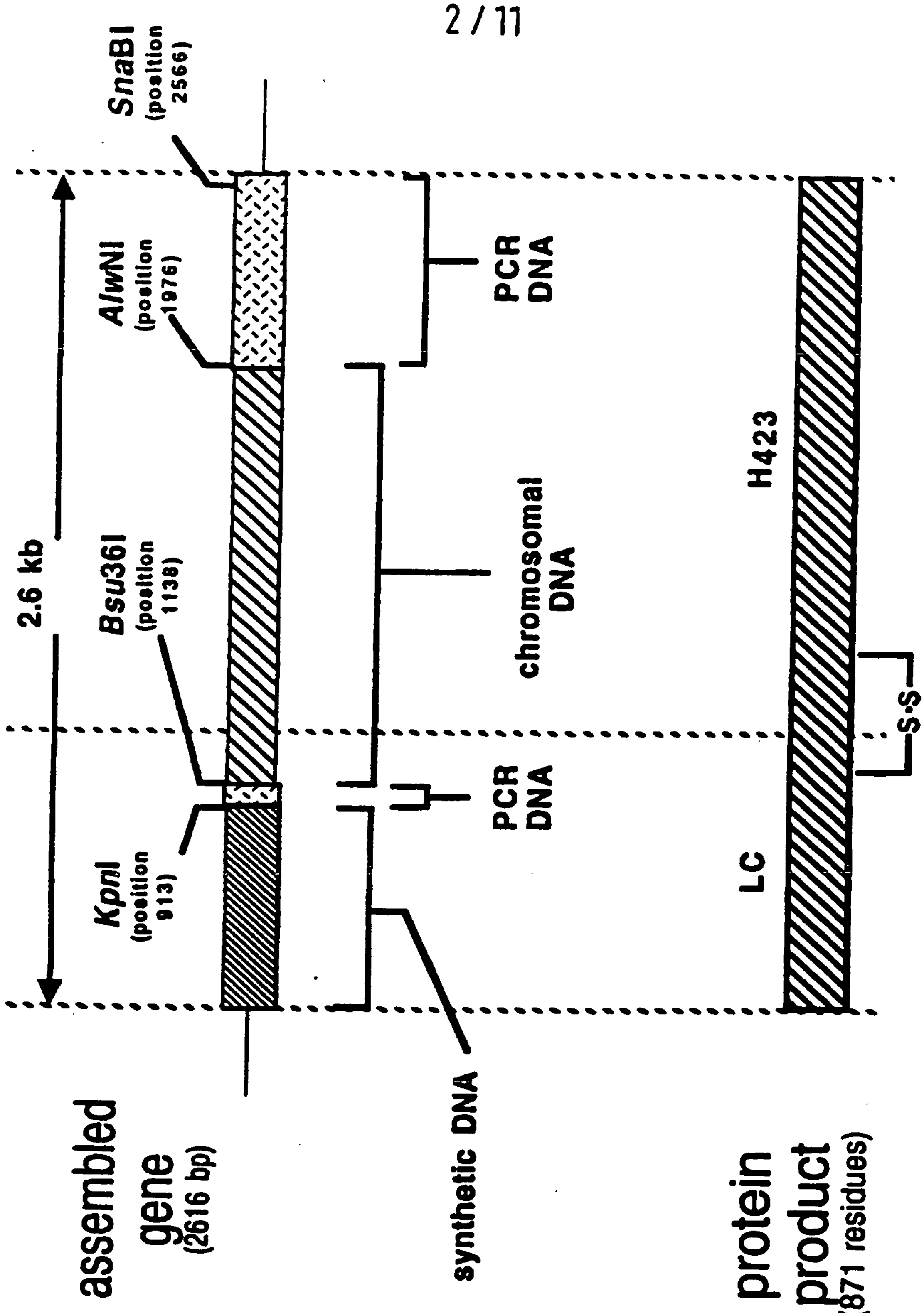

SnaBI (position 2566)
AlwNI (position 1976)
Bsu36I (position 1138)
KpnI (position 913)
2.6 kb
PCR DNA
chromosomal DNA
PCR DNA
synthetic DNA
H423
LC
S-S
assembled gene (2616 bp)
protein product (871 residues)

FIG. 2

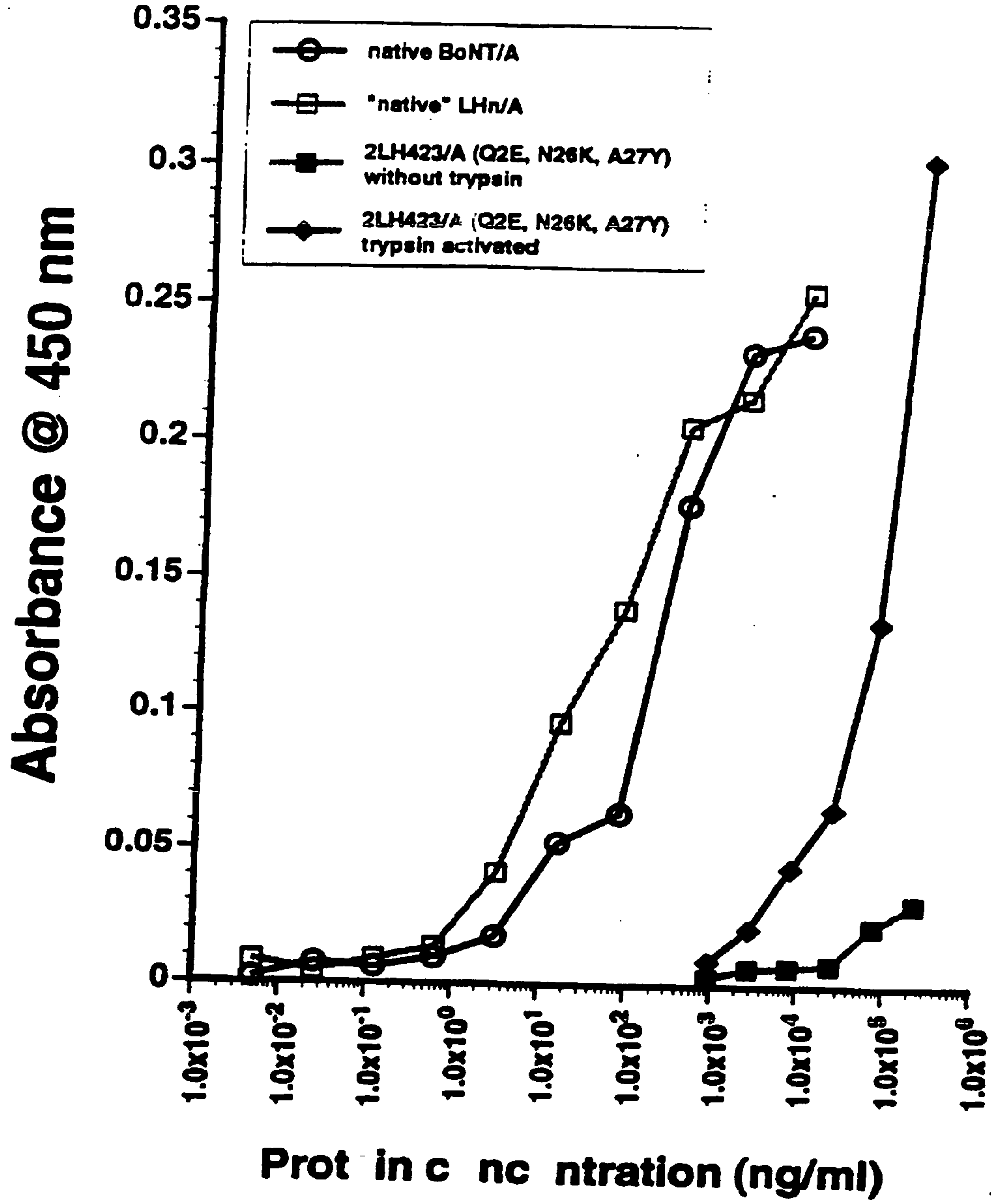

native BoNT/A
"native" LHn/A
2LH423/A (Q2E, N26K, A27Y) without trypsin
2LH423/A (Q2E, N26K, A27Y) trypsin activated
Absorbance @ 450 nm
0.35
0.3
0.25
0.2
0.15
0.1
0.05
0
1.0x10⁻³
1.0x10⁻²
1.0x10⁻¹
1.0x10⁰
1.0x10¹
1.0x10²
1.0x10³
1.0x10⁴
1.0x10⁵
1.0x10⁶
Prot in c nc ntration (ng/ml)

FIG. 3

LH_{423}/A	M Q F V N K Q F N Y K D P V N G V D I A Y I K I P N A G Q M Q P V (1 under M)	(Seq I.D. 2)
$_{23}LH_{423}/A$ (Q_2E, $N_{26}K$, $A_{27}Y$)	G S P G I H M T S T R L Q K L L E F E L P (1 under G) G T M E F V N K Q F N Y K D P V N G V D I A Y I K I P K Y G Q M Q P V (24 under M)	(Seq I.D. 4)
$_{2}LH_{423}/A$ (Q_2E, $N_{26}K$, $A_{27}Y$)	G S M E F V N K Q F N Y K D P V N G V D I A Y I K I P K Y G Q M Q P V (1 2 3 under G S M)	(Seq I.D. 6)
Native BoNT/A, *C. botulinum* 2169 Thompson *et al.* 1990	M Q F V N K Q F N Y K D P V N G V D I A Y I K I P N A G Q M Q P V (1 under M)	
Native BoNT/A, *C. botulinum* 62A Binz *et al.* 1990	M P F V N K Q F N Y K D P V N G V D I A Y I K I P N A G Q M Q P V (1 under M)	

$\square$ = REGIONS OF NON-IDENTITY WITH THE NATIVE SEQUENCES.

FIG. 4

```
1321/441                                  Eco47 III  1351/451
TCA TTA GAT AAA GGA TAC AAT AAG agc gct gat ggg GCA TTA AAT GAT TTA TGT ATC AAA
S   L   D   K   G   Y   N   K   S   A   D   G   A   L   N   D   L   C   I   K
                                                H-chain
```

FIG. 5

```
1321/441                                        1351/451
TCA TTA GAT AAA GGA TAC AAT AAG atc gaa ggt cgt tgc gat ggg GCA TTA AAT GAT TTA
S   L   D   K   G   Y   N   K   I   E   G   R   C   D   G   A   L   N   D   L
                                Factor Xa                   H-chain
                                protease motif
```

FIG. 6

```
2587/863                                2617/873                                IGF-1
TAC GTA GAT AAT CAA AGA TTA TTA TCT ACA TTT ACT GAA TAT ATT AAG TCT AGG CCT GGA
Y   V   D   N   Q   R   L   L   S   T   F   T   E   Y   I   K   S   R   P   G
2647/883                                2677/893
CCG GAG ACG CTC TGC GGG GCT GAG CTG GTG GAT GCT CTT CAG TTC GTG TGT GGA GAC AGG
P   E   T   L   C   G   A   E   L   V   D   A   L   Q   F   V   C   G   D   R
2707/903                                2737/913
GGC TTT TAT TTC AAC AAG CCC ACA GGG TAT GGC TCC AGC AGT CGG AGG GCG CCT CAG ACA
G   F   Y   F   N   K   P   T   G   Y   G   S   S   S   R   R   A   P   Q   T
2767/923                                2797/933
GGT ATC GTG GAT GAG TGC TGC TTC CGG AGC TGT GAT CTA AGG AGG CTG GAG ATG TAT TGC
G   I   V   D   E   C   C   F   R   S   C   D   L   R   R   L   E   M   Y   C
2827/943                                2857/953
GCA CCC CTC AAG CCT GCC AAG TCA GCT GAA GCT TAG
A   P   L   K   P   A   K   S   A   E   A   stop
```

FIG. 7

```
2587/863                                2617/873                                CtxA14
TAC GTA GAT AAT CAA AGA TTA TTA TCT ACA TTT ACT GAA TAT ATT AAG TCT AGG CCT CAA
Y   V   D   N   Q   R   L   L   S   T   F   T   E   Y   I   K   S   R   P   Q
2647/883                                2677/893
TCT AAA GTT AAA AGA CAA ATA TTT TCA GGC TAT CAA TCT GAT ATT GAT ACA CAT AAT AGA
S   K   V   K   R   Q   I   F   S   G   Y   Q   S   D   I   D   T   H   N   R
2707/903
ATT AAG GAT GAA TTA TGA
I   K   D   E   L   stop
```

FIG. 8

```
2587/863                                          2617/873
TAC GTA GAT AAT CAA AGA TTA TTA TCT ACA TTT ACT GAA TAT ATT AAG TCA GGC CTG AAT
Y   V   D   N   Q   R   L   L   S   T   F   T   E   Y   I   K   S   G   L   N
2647/883                                          2677/893
TCC CCG GGT GCA GCT CAT TAT GCG CAA CAC GAT GAA GCC GTA GAC AAC AAA TTC AAC AAA
S   P   G   A   A   H   Y   A   Q   H   D   E   A   V   D   N   K   F   N   K
2707/903                                          2737/913
GAA CAA CAA AAC GCG TTC TAT GAG ATC TTA CAT TTA CCT AAC TTA AAC GAA GAA CAA CGA
E   Q   Q   N   A   F   Y   E   I   L   H   L   P   N   L   N   E   E   Q   R
2767/923                                          2797/933
AAC GCC TTC ATC CAA AGT TTA AAA GAT GAC CCA AGC CAA AGC GCT AAC CTT TTA GCA GAA
N   A   F   I   Q   S   L   K   D   D   P   S   Q   S   A   N   L   L   A   E
2827/943                                          2857/953
GCT AAA AAG CTA AAT GAT GCT CAG GCG CCG AAA GTA GAC AAC AAA TTC AAC AAA GAA CAA
A   K   K   L   N   D   A   Q   A   P   K   V   D   N   K   F   N   K   E   Q
2887/963                                          2917/973
CAA AAC GCG TTC TAT GAG ATC TTA CAT TTA CCT AAC TTA AAC GAA GAA CAA CGA AAC GCC
Q   N   A   F   Y   E   I   L   H   L   P   N   L   N   E   E   Q   R   N   A
2947/983                                          2977/993
TTC ATC CAA AGT TTA AAA GAT GAC CCA AGC CAA AGC GCT AAC CTT TTA GCA GAA GCT AAA
F   I   Q   S   L   K   D   D   P   S   Q   S   A   N   L   L   A   E   A   K
3007/1003                                         3037/1013
AAG CTA AAT GAT GCT CAG GCG CCG AAA GTA GAC TAG
K   L   N   D   A   Q   A   P   K   V   D   *
```

FIG. 9

LH423/A

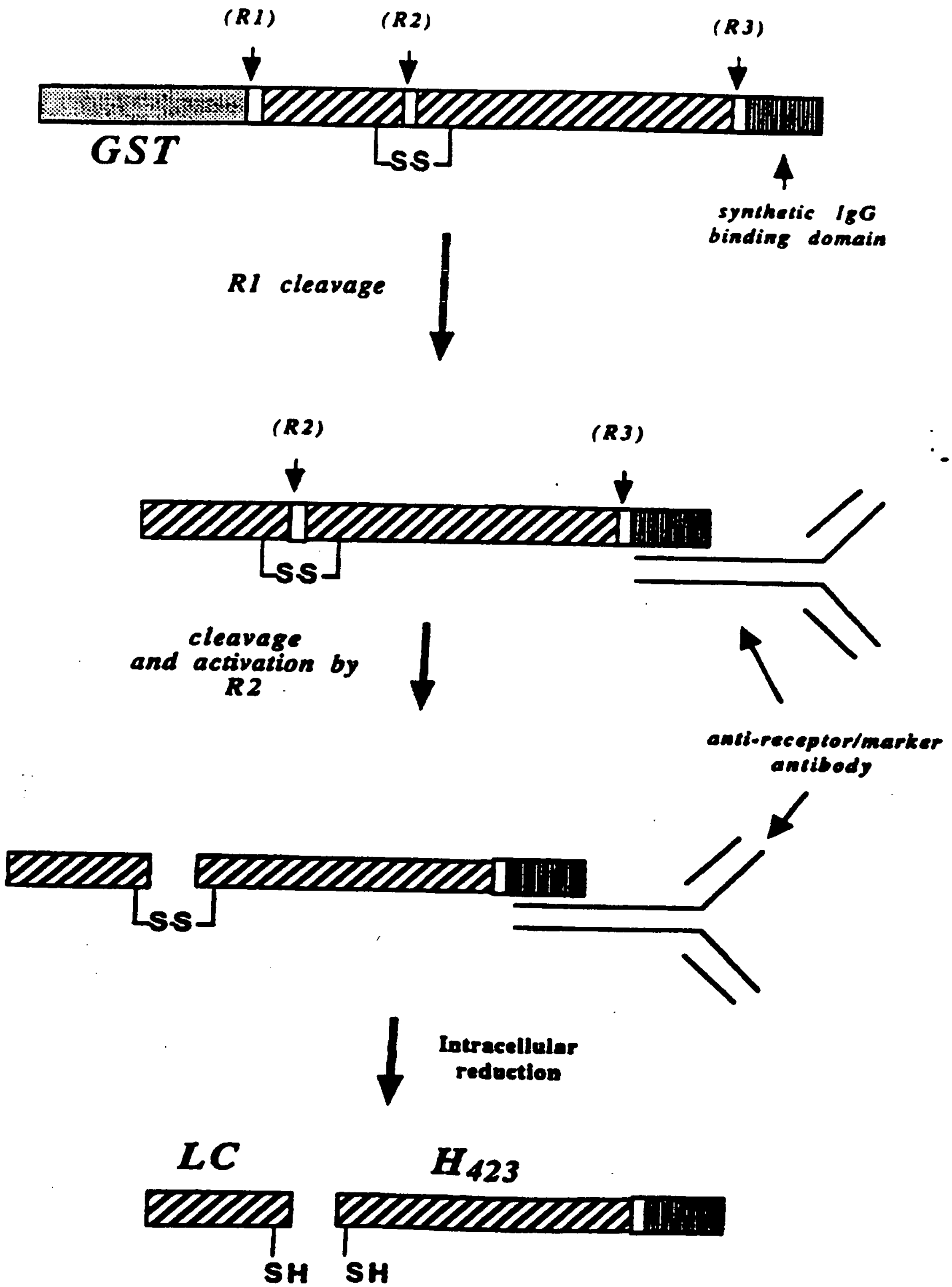


FIG. 10

LH_{423}/A

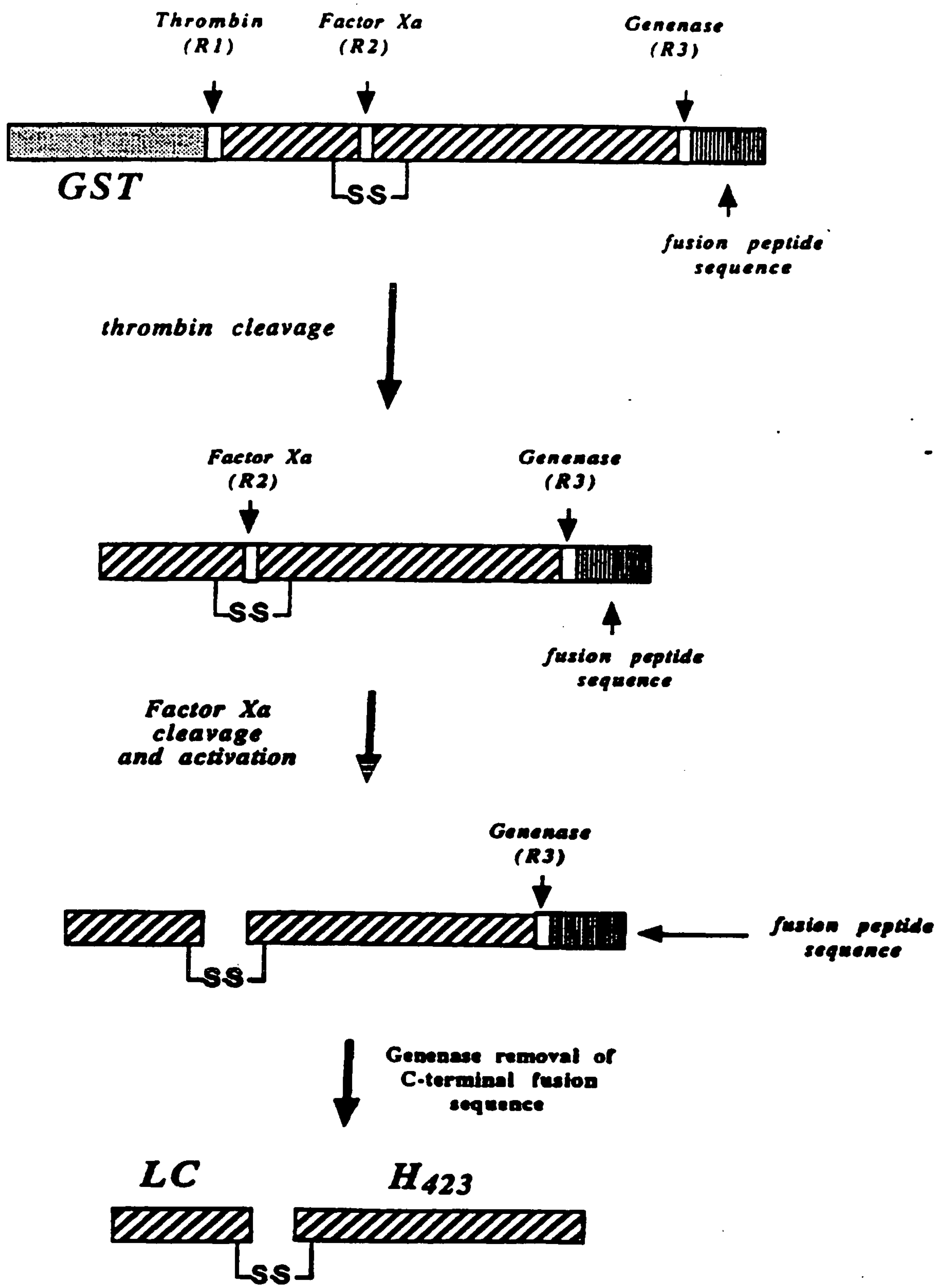


FIG. 11

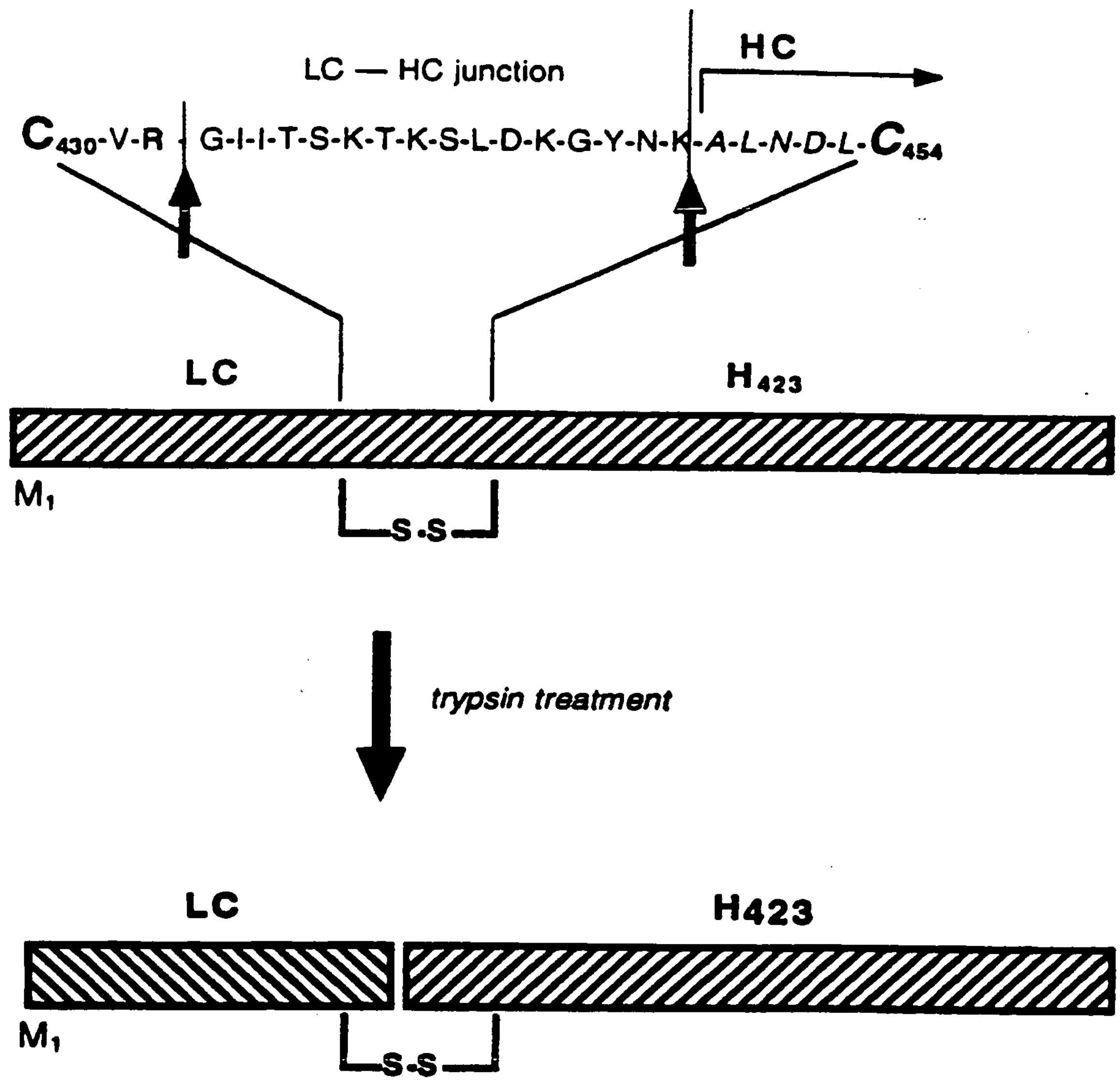

HC
LC — HC junction
C430-V-R - G-I-I-T-S-K-T-K-S-L-D-K-G-Y-N-K-A-L-N-D-L-C454
LC
H423
M1
S-S
trypsin treatment
LC
H423
M1
S-S

FIG. 12

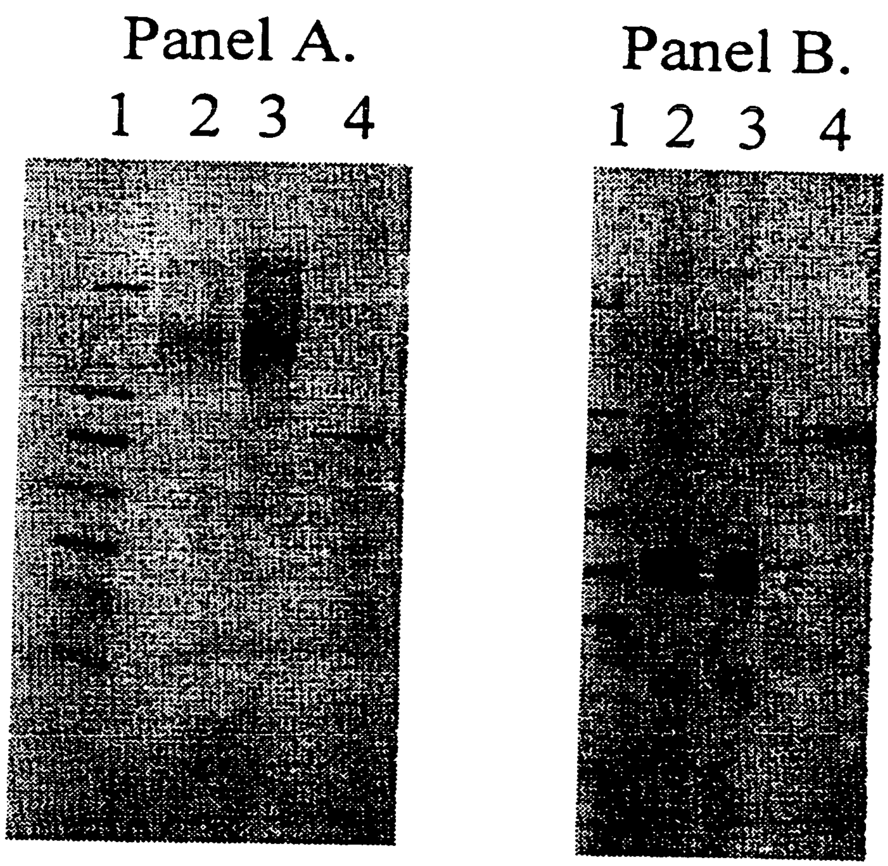


FIG. 13

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 97/02273

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/31 C12N1/21 C12P21/02 C07K14/33 A61K38/16
A61K39/08

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12P A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WO 96 12802 A (OPHIDIAN PHARM INC ;WILLIAMS JAMES A (US); PADHYE NISHA V (US); KI) 2 May 1996 see the whole document ---	1-52
X	KURAZONO H ET AL: "Minimal essential *domains* specifying toxicity of the *light* *chains* of tetanus toxin and botulinum neurotoxin type A." J BIOL CHEM, JUL 25 1992, 267 (21) P14721-9, UNITED STATES, XP002047910 see table II ---	1-52
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
9 December 1997	30.01.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hillenbrand, G

INTERNATIONAL SEARCH REPORT

Inten .onal Application No
PCT/GB 97/02273

C.(Continuation) DOCUMENTS C NSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	LI ET AL: "A SINGLE MUTATION IN THE RECOMBINANT LIGHT CHAIN OF TETANUS TOXIN ABOLISHES ITS PROTEOLYTIC ACTIVITY AND REMOVES THE TOXICITY SEEN AFTER RECONSTITUTION WITH NATIVE HEAVY CHAIN" BIOCHEMISTRY, vol. 33, no. 22, 1994, pages 7014-7020, XP002015938 see the whole document ---	1
A	BINZ T ET AL: "THE COMPLETE SEQUENCE OF BOTULINUM NEUROTOXIN TYPE A AND COMPARISON WITH OTHER CLOSTRIDIAL NEUROTOXINS" JOURNAL OF BIOLOGICAL CHEMISTRY, vol. 265, no. 16, 5 June 1990, pages 9153-9158, XP002009348 see the whole document -----	1,26,35

INTERNATIONAL SEARCH REPORT

Inter. ...onal Application No
PCT/GB 97/02273

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9612802 A	02-05-96	US 5601823 A	11-02-97
		US 5196193 A	23-03-93
		AU 3968395 A	15-05-96
		EP 0796326 A	24-09-97
		FI 971732 A	23-06-97
		NO 971868 A	24-06-97
		PL 320214 A	15-09-97
		ZA 9508990 A	15-05-96
		AU 6653894 A	04-07-94
		CA 2150935 A	23-06-94
		EP 0671902 A	20-09-95
		WO 9413264 A	23-06-94
		US 5466672 A	14-11-95
		US 5599539 A	04-02-97
		AU 638786 B	08-07-93
		AU 6895191 A	31-05-91
		EP 0498854 A	19-08-92
		WO 9106306 A	16-05-91
		US 5443976 A	22-08-95
		US 5340923 A	23-08-94